

ABSTRAK

Manajemen persediaan barang di gudang kerap mengalami ketidaksesuaian jumlah kardus akibat pencatatan manual yang tidak akurat, yang dapat mengganggu efisiensi operasional dan akurasi data stok. Penelitian ini mengembangkan sistem deteksi kardus berbasis *Mask R-CNN* dengan dua konfigurasi *backbone*: *ResNet-50* (default) dan *MobileNetV3 Small* (custom). Tujuan utama meliputi implementasi deteksi kardus, penentuan konfigurasi *hyperparameter* optimal untuk *MobileNetV3 Small*, serta perbandingan akurasi, ukuran model, dan kecepatan inferensi kedua model. *Dataset* dikumpulkan dari empat kategori kardus dan dianotasi menggunakan *Roboflow*, lalu dilatih dengan metode *Stochastic Gradient Descent*. Model custom menggunakan konfigurasi optimal dengan *learning rate* 0,002, *batch size* 2, dan 20.000 iterasi. Hasil menunjukkan model custom lebih efisien, dengan pengurangan parameter dari 43,9 juta menjadi 20,8 juta, serta peningkatan kecepatan inferensi hingga 8,66 *FPS*. Walaupun model default unggul dalam akurasi objek kecil, model custom tetap kompetitif dan layak untuk aplikasi *real-time* di perangkat terbatas.

Kata kunci: *Mask R-CNN*, *MobileNetV3 Small*, Deteksi Kardus, Konfigurasi Model



ABSTRACT

Inventory management in warehouses often experiences discrepancies in the number of cardboard boxes due to inaccurate manual recording, which can interfere with operational efficiency and stock data accuracy. This research develops a Mask R-CNN based cardboard detection system with two backbone configurations: ResNet-50 (default) and MobileNetV3 Small (custom). The main objectives include implementation of cardboard detection, determination of the optimal hyperparameter configuration for MobileNetV3 Small, and comparison of accuracy, model size, and inference speed of both models. Datasets were collected from four cardboard categories and annotated using Roboflow, then trained with the Stochastic Gradient Descent method. The custom model used the optimal configuration with a learning rate of 0.002, batch size 2, and 20,000 iterations. Results show that the custom model is more efficient, with a reduction in parameters from 43.9 million to 20.8 million, as well as an increase in inference speed to 8,66 FPS. While the default model excels in small object accuracy, the custom model remains competitive and feasible for real-time applications on limited devices.

Keywords: Mask R-CNN, MobileNetV3 Small, Cardboard Detection, Model Configuration

